

---

| RESEARCH ARTICLE

## Meaning Construction of Multimodal Synergy in Documentary Discourse: Taking *The Lockdown: One Month in Wuhan* as an Example

Feifei Zhang

Tiangong University, Tianjin, China

Corresponding Author: Feifei ZhangTi, E-mail: 707326859@qq.com

---

| ABSTRACT

This article, based on Zhang Delu's synthetic theoretical framework for multimodal discourse analysis (MDA), conducts a multimodal discourse analysis on the documentary-*The Lockdown: One Month in Wuhan*. It is found that the documentary discourse involves two necessary modalities: language modality and image modality. The main relationship between them is complementary reinforcement. The common relationships between necessary modality and selective modality are interaction, primary and secondary, and context interaction. Although the overlapping relationship has little effect on the meaning construction of the discourse, its existence is justified in some cases. Through the analysis of the synergy of different modes, it is hoped that the article will provide some enlightenment for multimodal relationships in the meaning construction of documentary discourse.

| KEYWORDS

Multimodality, multimodal discourse analysis, mode synergy, relationship, meaning construction

| ARTICLE DOI: 10.32996/ijllt.2022.5.6.7

---

### 1. Introduction

The development of modern media technology has greatly changed communication channels and ways of information transfer. Various media use images, animation, music, font size, color, spatial layout, and so on to communicate. As discourse becomes increasingly complex and discourse analysis theory develops, people have realized that meaning is made jointly by all the semiotic resources in discourse, including non-language resources such as action, image, sound, expressions, gestures, diagrams, pictures, etc. These semiotic resources are called modes or modalities. The phenomenon of using auditory, visual, tactile, and other senses to communicate through language, image, sound, action, and other means and symbolic resources is the multimodal analysis (Zhang Delu, 2009). In this sense, it is not sufficient to provide a full account of meaning in social situations only by language, for it provides the only "partial meaning". It is necessary to fundamentally explore different semiotic categories and the forms, meaning systems, functions, and application scope of different modalities, as well as the cooperation between modal resources in meaning-making.

Scholars of systemic functional semiotics began to study the synergy of different modalities in the 1990s, and text-image relation is one hot field. R. Barthes (1997) proposed three relationships-anchorage, illustration, and relay between text and image, in his essay *Rhetoric of the image*, published in 1977. Martinec and Salway (2005) make use of the logico-semantic relations developed by Halliday. They proposed that status of image and text includes equal and unequal relation, and logico-semantic relation falls into the categories of extension and projection. Terry Royce (1998) explores inter-semiotic complementarity in page-based multimodal text from ideational, interpersonal, and textual perspectives.

In recent years, more and more attention has been paid to the synergy of different modes among domestic scholars. (Li Zhanzi 2003; Hu Zhuanglin 2006; Zhu Yongsheng 2007). Zhang Delu (2009) sets up a synthetic framework for multimodal discourse analysis, in which he proposes the complementary relationship and non-complementary relationships among different modalities.

**Copyright:** © 2022 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (<https://creativecommons.org/licenses/by/4.0/>). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

Based on that, Zhang Delu and Yuan Yanyan (2011) investigate the synergy of different modes by taking TV weather forecasts as an example and find some principles underlying the synergy between different modalities. Within the framework of system functional grammar (SFG), Chen Songjing (2016) analyzes stage narratives in an English recitation contest and explores how narrative modality interaction influences the meaning by using video analytics software MMAVideo2.0, which makes up for the deficiency of the previous research in quantitative analysis. Geng Jingbei and Chen Zijuan (2014) investigate the dynamic multimodal discourse features of documentaries from discourse context, discourse construction, and discourse cohesion and find that context's potential guide to a generation of discourse, the textual function of various symbols, as well as discourse cohesion means play important roles in the dynamic multimodal discourse construction of documentaries.

Recent years have witnessed an increasing study on dynamic multimodal discourse, such as films, classroom discourses, promotional videos, and stage narratives, but the studies on documentary's multimodal synergy are very limited, especially the documentary, which plays a key role in the construction of China national discourse system.

## 2 Theoretical Framework of Dynamic Multimodal Discourse Analysis

Zhang Delu set up a synthetic theoretical framework for MDA based on Halliday's theory. This framework consists of four levels: (1) the level of context of culture; (2) the level of context; (3) the level of content; (4) the level of expression. It is designed and borrowed from five levels in SFL theory. That is why we say SFL is the most applicable theoretical framework to the MDA.

The level of culture consists of people's thinking model, philosophy, habits, and customs and all of the unspoken rules, and the communicative procedure or structure potential which can realize the ideology. The context of the situation does not work without it. At the level of context, communication is subject to contextual factors, including field, tenor, and mode, which can determine contextual factors. At the level of content, ideational meaning, interpersonal meaning, and textual meaning are determined by field, tenor, and mode. At the level of form, the discourse meaning is realized by the interaction of different modes. Each mode has its own formal system, such as visual grammar, auditory grammar, tactile grammar, etc. The relations between different lexicogrammar of different modalities can be defined into two kinds: complementary and non-complementary (cf. Table 1).

Table 1: Layout of modal relationship of Zhang Delu (2019)

Relation Types	Main Relation	Concrete Relation
Complementary	Reinforcement	Highlighting
		Primary and secondary
		Extension
	Non-reinforcement	Interaction
		Coalition
		Coordination
Non-complementary	Overlapping	Redundancy
		Exclusion
		Counteraction
	Inclusion	Whole and part
		Abstract and concrete
	Context Interaction	Context-independence
		Context-dependence

The article illustrates the cultural and situational context of the documentary and mainly explains the synergy of modalities in the form lane.

## 3 Data and Research Methods

### 3.1 Data

The article studies a documentary called *The Lockdown: One Month in Wuhan*. It is an English news documentary produced by CGTN, lasting for 33 minutes and 14 seconds. The documentary is dedicated to all those who have been battling tirelessly against the COVID-19 virus. By combining the time nodes of important events, the first-hand news of Wuhan is presented to overseas viewers. It was broadcast on CGTN on February 28, 2020, and also delivered at the same time on the CGTN official website, CGTN application, and CGTN overseas social platform accounts (YouTube, Twitter, Facebook). Once it was released, it attracted great attention from overseas media. More than 21 countries and 165 overseas TV channels broadcast it in succession. In this sense, it is of great value to study.

### **3.2 Research methods**

A qualitative approach with a case study is adopted in the article, with the quantitative approach as a supplement. Firstly, the author defines and classifies different modalities used in the documentary, and then a video analysis software MMAVideo2.0 is used to transcribe the video. The main step is to classify and label the important modal relationships appearing in the video and then count the distributional proportion of these relationships in the total length of the video. In order to explain the detailed relationships clearly, the author picks up six shots to probe into the role that the involved modalities play. Finally, the article explores how mode synergy affects the meaning making of documentary multimodal discourse and provides enlightenment to national discourse image building.

## **4 Contextual Analysis of Documentary Meaning Construction**

### **4.1 Context of culture**

Halliday regards language as a social semiotic that depends on context. The meaning and function of language in practice are influenced and restricted by specific social and cultural factors. To some extent, the choice of modalities in discourse also depends on social and cultural context. The cultural context is the history, culture, customs, way of thinking, moral ideas, and values of the speech community in which the speaker is located, and people who belong to the speech community generally understand the meaning in the context. It means that people adopt different forms of communication according to the different purposes of communication.

The dynamic multimodal discourse-*The Lockdown: One Month in Wuhan*, aims at enhancing national cohesion and shaping a positive national image around the world. CGTN uses bilingual subtitles to display a real virus-fighting situation in China worldwide. As we can see in the documentary, China has adopted "extraordinary measures" to tackle the COVID-19 epidemic over the past months and actively engaged in international cooperation in curbing the spread of the virus, which contributes to the building of a community with a shared future for mankind.

### **4.2 Context of situation**

According to the theory of SFL, context of situation is the concrete embodiment of the context of culture, including three variables: field, the social activities involved in discourse; tenor, the social role between two participants; mode, the medium, and channel of communication. In the documentary, it can be described specifically as follows.

Field: The development of COVID-19 in Wuhan from January 1, 2020, to February 24, 2020.

Tenor: China International Television, worldwide documentary viewers, CGTN reporters, and the general public.

Mode: CGTN makes use of audio-video media to present the dynamic development of Covid-19 fighting in Wuhan; the reporters try to use Mandarin Chinese to talk with people of all works of life in Wuhan, and the discourse designer chooses English as the main auditory modality for the background commentary to convey sound meaning; dynamic picture formed by the continuous change of body media (gaze, expression, gesture, and other body languages) and non-body media(background, bilingual subtitles) is used as a visual modality to spread the image of Wuhan's anti-epidemic process.

## **5 Mode Synergy in Documentary Multimodal Discourse**

### **5.1 Modal forms**

The documentary narrative contains a variety of mode resources. The author first shows a multimodal representation and sorts out different modalities in this dynamic multimodal discourse. There are two main modalities in the documentary (cf. Table 2): auditory modality and visual modality, among which the oral language of auditory modality and the written language of visual modality can be regarded as language modalities.

Table 2: Modal forms in *The lockdown: One Month in Wuhan*

Modal Forms	Detailed Modal Elements	
Auditive modalities	Special effect sound	
	Background music	
	Oral language	Spoken narrative
Visual modalities	Written language	Other spoken languages
		Subtitle
	Image	Written words
		Filmic image
	Action	Static picture
		Gesture
	Expression	

### 5.2 Relationships between necessary modalities

Visual and auditory elements are the two indispensable parts of the documentary. Visual modality is mainly embodied in the subtitle and filmic image, which not only successfully expresses the time and place of discourse but also expresses the transfer of time and space through lens movement and switching. Auditive modality is primarily represented in oral language. The ambiguity of meaning expressed by image and the precision of meaning produced by language match each other well and finally achieve the best communication effect. Language modality and image modality are common and basically throughout the documentary; thus, the author regards these two modalities as necessary modalities and others are optional modalities.

It is found that the relationships between necessary modalities in this dynamic discourse have three main types: complementary reinforcement, non-reinforcement coordination, and abstract and concrete. The time duration and ratio of these relationships calculated by the software MMAVideo 2.0 are shown in Table 3.

Table 3: The distribution of relation types of necessary modality collaboration

Modality Collaboration	Relation Type	Concrete representation of mode synergy	Time Duration	Ratio
language modality + Image modality	Complementary Reinforcement	Spoken narrative/subtitle + dynamic image	1004s	50.33%
	Complementary Non-reinforcement Coordination	spoken language + speaker image	701s	35.14%
	Non-complementary Inclusion Abstract and concrete	Oral language + static picture/dynamic image	148s	7.4%
Total			1853s	92.87%

According to Table 3, the main relationship between language modality and image modality is complementary reinforcement, whose time duration makes up roughly 50.33% of the whole length of the video. The second common relationship is non-reinforcement coordination, constituting 35.14%, which mainly represents the synergy of the speaker's oral language and the speaker's image. The two modalities also form the relationship of non-complement inclusion, in which the image modality provides the concrete image of what oral language talks about. The proportion of this relationship accounts for 7.4% of the total time. In total, the synergy of language modality and image modality almost make up 92.87% of the whole video, and the rest ratio was shown in the match of other modes, like music and image.

#### 5.2.1 Complementary reinforcement

If the discourse of one mode cannot fully express its meaning or cannot express its full meaning, and needs to be supplemented by another mode, then the relationship between modes is called complementary relation. It can be further classified into reinforcement relation and non-reinforcement relation. A reinforcement relation is one in which one mode is the dominant form of communication, and another or more forms help reinforce it. It involves highlighting, primary and secondary, and extension.

Highlighting means that one modality provides the background for another and makes it the foreground. In Figure 1, video footage acts as a background, showing the ambulance speeding the patient away from the scene. At the same time, dynamic digital number popping up reflects a rapid increase in deaths. The dynamic image provides the background for the written words and makes them stand out to catch viewers' attention. By comparing the death number between SARS and COVID-19, combined with the background high-speed ambulance, mode cooperation between image and words highlights the severity of the new epidemic on threatening life.



Figure 1: written language pop up

If one modality expands discourse meaning with the help of another one, then these two modes are in an extension relation. During an interview with a community worker, the worker's child said that her mom belonged to Wang Wang Team. Some foreigners do not understand the meaning of the Wang Wang Team. Here, a short video clip followed immediately to show that this address is from popular animation, *PAW Patrol* in China, which tells stories about the dog patrol's kinds acts. The short clip indirectly strengthens the dedicative image of volunteers in people's hearts, which enriches the meaning of spoken language.



Figure 2: a video clip from *PAW Patrol*

A short video shot by a person was cited in the documentary, which shows a huge crowd gathered in a local hospital. With the video maker commending, "I have never seen anything like that". Dynamic image (the primary mode), combined with oral language (the secondary mode), accurately demonstrates the overwhelming demand for medical resources and the reality that the hospital was overloaded.

### **5.2.2 Non-reinforcement coordination**

The synergy between the speaker's image and the speaker's language is a typical relation of coordination. While showing the speaker's image, his or her oral language should be played simultaneously. Otherwise, communication will be difficult to understand. Image and sound are equally important in this sense. Spoken language collected in the documentary includes the official speakers' speeches, doctor-patient dialogues, interview communication, and other oral languages except for spoken narrative. The distribution of spoken language in the documentary accounts for 701 seconds (35.14%).

### **5.2.3 Abstract and concrete**

The relationship between abstract and concrete means that the second modality does not provide new meaning, but it adds more concrete and vivid information to the first modality's illustration. In the documentary, the doctor tells us that when he takes off his protective gear, he will feel itchy. At this time, several pictures with rashes all over the body are presented. In this sense, oral language and pictural image are in an abstract-concrete relation because pictures show a concrete situation of rash in his body and powerfully express the difficulties he has suffered. Besides, when it comes that the origins of the virus may come from wild animals, pictures of those animals are presented. It can also be considered as an abstract and concrete relationship between language and image.



Figure 3: rush in body



Figure 4: wild animal

### 5.3 Relations between necessary modalities and selective modalities

Modality relationships in documentary discourse are complex and changeable, but the relationships between necessary modalities and selective modalities are relatively stable. The author finds three common relationships in the documentary that is an intersection, primary and secondary, and context interaction.

#### 5.3.1 Intersection

Intersection relationship means that two modalities are used alternatively and simultaneously to express complete meaning. In Figure 5, the doctor takes off three layers of protective gloves and explains the process of how to do it at the same time. It explains what he or she is doing as he or she does it. In this regard, two modalities (gesture and oral language), not opposing each other in meaning expressing, are in an intersection relationship.



Figure 5: pick off the gloves

#### 5.3.2 Primary and secondary

In this relationship, one modality is primary in articulating discourse meaning. Another is secondary to assist it. The director uses a close-up to show the mixed emotions of female nurses and other patient members outside the isolation ward when they talk about the difficulties of isolation and healthcare needs. Their eyebrows, lips, and twitching noises convey virtually every subtle nuance of sorrow and grief. Language modality is primary, and the visually intriguing image is secondary, assisting in expressing the bitterness of suffering families.

#### 5.3.3 Context interaction

The relationship between background music and the image is context independent in the whole discourse. It finds that background music runs through the documentary, with the volume constantly changing. When there is a commentary language, background music is in the background position. When there is no spoken language, it is prominent in the foreground position. In this documentary, background music is a kind of soothing music playing in a loop with a simple melody and steady rhythm. The context or atmosphere created by the music has nothing to do with the overall meaning construction, neither lifting the spirit to fight against the coronavirus nor giving hope to the future.

In contrast, the context created by special effect sound is directly involved in meaning making. The special sound modality forms the relationship of context dependence with the visual modality. The special effect sounds collected in the discourse (3%) include the police car and ambulance sirens, community horns, medical equipment sound, mask-manufacturing sound, door knocks, emergency telephone rings, etc. They directly participate in the construction of the overall meaning, reflecting the urgency of saving lives and the efficiency of community work and the work resumption.

### 5.4 Relationships between modes within the same modality

Modalities enjoy various representatives. Detailed modal elements (spoken language and written language) within the same modality (language modality) can be used concurrently in meaning making, sometimes showing overlapping relationships. In Zhang Delu's synthetic theoretical framework for MDA, the relation of overlapping, including exclusion, counteraction, and redundancy, should be avoided in meaning construction because two or more modalities in these relationships could damage and detract meaning conveyed by each other. However, the author finds it does not mean that the simultaneous use of two modes is wasteful as long as they are in an overlapping relationship. Counteraction relationships and redundancy relationships are allowed to exist for special reasons in some circumstances in the documentary discourse.

### 5.4.1 Redundancy

In terms of language modalities, subtitles and spoken narrative are two commonly used modes in documentary discourse. They appear at the same time, expressing the same information rather than adding or subtracting new or old information. Generally, the spoken narrative is called "overt translation," and the subtitle is "covert translation". And whether to read subtitles or listen to narration depends on a variety of factors such as age, educational level, reading habits, the conception of the audience and national language policies, and so on. Besides, CGTN's media property and the context of the situation of the documentary require bilingual subtitles, which meet the demand of both domestic viewers and foreign audiences. In this sense, the redundancy relation is justified.

Similarly, in Figure 6, the textual words in the picture also appear in the form of narration. Although certain words or sentence structures are replaced, the content of information is still the same. That means the meaning made by these two modes is exactly the same; that is, China's President Xi Jinping stresses that the people's safety and health are the priority. The meaning represented by two modal forms does not complement or enforce each other. It is a redundancy relationship. The function of the usage of two modes, on the one hand, conforms to the development process of narration and provides convenience for the hearer to read; on the other hand, political discourse emphasizes accuracy, and the interpretation of information cannot be changed greatly, so repeated emphasis demonstrated by written words and oral words appears simultaneously.



Figure 6: China's President Xi Jinping's work construction

### 5.4.2 Counteraction

Counteraction means the existence of one mode or modality destroys or interferes with the meaning that another expresses. Within visual modalities, it is found that the icon image of CGTN is always in the upper left corner of the documentary, which interferes with the visual dynamic image as a whole. However, because of the copyright attributes, the icon can be acceptable, especially for an international propaganda video. It is just the size of the icon that matters. The documentary maker has tried to avoid the impact of icon size on meaning interpretation, so it is designed as a small one on the upper left of the whole composition.

## 6. Discussion

### 6.1 The construction of national image by documentary mode synergy

The basic view of sociolinguistics holds that the essence of language is a kind of social practice. As the epidemic is still raging globally, discourse construction around COVID-19 is affected by factors such as geopolitics, history, culture, and national interests, showing complex and changeable characteristics. Thus, how China's national discourse expresses this event, which affects the world's human destiny, is bound to be embedded in the broader narrative of "human community with a shared future". Under the background of globalization, the key to enhancing the influence of national discourse is to enhance the narrative ability of the image of the contemporary country, its society, and its people. It is vital to reflect the real situation of Chinese society through narrative creation and provide concepts that can be understood, experienced, and shared by the world.

In the past decades, the study of MDA has developed rapidly, and the application of relevant theories in the field of media has attracted increasing attention. However, there are few types of research on the discourse construction of national image from the perspective of multimodal discourse. CGTN is the main body of national discourse communication, and the documentary *The Lockdown: One Month in Wuhan*, which reflects the lockdown situation in Wuhan, is the national discourse construction in this major global security and health event. In *The Lockdown: One Month in Wuhan*, a real China has been made by using different modalities and a collection of different modes.

For example, the documentary uses chatting record screenshots of WeChat, video clips of TikTok, and an official video clip from the World Health Organization press conference. These materials constitute the information source of the documentary, ensuring the credibility and authority of the narrative to a certain extent.

When the documentary shows the building process of Huoshenshan Hospital, the text mode on the upper right represents the dynamic change of dates from January 27 to 29. Hospital construction lives under the online supervisor, equipped with the date changing, showing the commendable speed of China's infrastructure capacity. If it were shown just with pictures, we might not

imagine a hospital could be built in just several days. Time goes by, and China's capacity for immediate response to disease displays the ability of resource integration, coordination, and management of the Chinese government, workers, and enterprise managers, as well as their selfless dedication and collectivism spirit.

The documentary also records the scene of medical teams from other provinces who are taking a collective oath after they arrive in Wu Han. They raised their right fist and swore loudly. The combination of gesture mode and language mode strongly evokes a solemn emotional experience of giving and mission. A narration of the event, accompanied by photos of the scene, will not evoke a strong sense like this.

Besides, in the documentary, reporters interview medical staff, military police, community members, experts, and party members. And the interviewee's discourse is the main communication message. As auxiliary means, text annotation, clothing, and party emblem help the audience understand respondents' identities, which also represses powerful Chinese cohesion and the determination of their resistance to the disease under the leadership of the Communist Party of China.

### **6.2 Implications for the use of modality in documentary**

Multiple modalities are used separately or simultaneously to express the meaning of discourse, serving different communicative purposes. For the documentary producers, they must reasonably design the synergy of different modalities according to the special context and situation.

In the construction of international communication, discourse designers should choose multimodal symbols with national characteristics based on the foreign audience's comprehension of discourse. And they should use all kinds of multimodal resources effectively to enhance the documentary expression effect, such as bilingual subtitles or English captions to enhance the overseas audience's understanding of the discourse's meaning.

The synergy of modalities should be based on the principle of positive effects to achieve the ultimate goal of communication, which means the selection and cooperation of modalities is of great importance. However, this does not mean that counteraction and redundancy relationships should be completely abandoned in the dynamic multimodal discourse design. If necessary, the size and frequency of a certain icon, for example, can be appropriately adjusted to participate in the discourse without damaging its meaning to the whole. But exclusion relation between modalities must be abandoned in a successful dynamic discourse; otherwise, the validity of the information will be questioned.

## **7. Conclusions**

The purpose of studying documentary discourse from the perspective of MDA is to help documentary makers to explain and evaluate the discourse construction so that they can have a systematic theoretical framework to discuss the important factors affecting the narrative effect of the documentary and have a scientific way to analyze and compare the differences of documentary narrative works.

This research has investigated the mode types and the synergy of different modalities in a documentary-The lockdown: One Month in Wuhan. The conclusion is that two necessary modalities have various representations, and the cooperation of the modes is changeable to achieve the clearest and most understandable communicative intent in different stages. MMAVideo2.0 is used to note the relationships among the modes, which makes up for the shortcomings of previous studies in quantitative analysis. But the limitation is that there is subjectivity in the classification and annotation of all kinds of relationships. It will be further studied in the future. The national discourse narrative in the post-epidemic era deserves more attention. It is hoped that the article can provide references for meaning construction in MDA based on the synergy of different modes and enlightenment for multimodality usage in the national discourse.

**Funding:** This research received no external funding.

**Conflict of Interest:** The authors declare no conflict of interest.

## **References**

- [1] Barthes, R. (1997). *"Rhetoric of the image" image-music-text*. London: Fontana.
- [2] Chen S. (2016). A Multimodal Discourse Analysis of Stage Narratives: Evidence from the Use of Stage Language in an English Recitation



- Contest. *Foreign Languages and Their Teaching*, (04):70-78+149. doi:10.13458/j.cnki.Flatt.004278.
- [3] Geng J, and Chen Z. (2014). The Dynamic Multimodal Discourse Analysis of Documentaries: Taking *Changsha, to the tune of Spring in a Pleasure Garden* ② from *Pointing to Stream and Hill: The Stories about Mao Zedong's Poetry* as an Example[J]. *Journal of Xi'an International Studies University*, 2014,22(04):24-28.doi:10.16362/j.cnki.cn61-1457/h.2014.04.007.
- [4] Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London and Melbourne: Arnold.
- [5] Hu Z and Dong J. (2006). How Is Meaning Construed Multimodally—A case study of a PowerPoint presentation contest. *Technology Enhanced Foreign Language Education*.03:3-12. DOI:CNKI:SUN: WYDH.0.2006-03-001.
- [6] Kress, G & T. van Leeuwen. (1996). *Reading Images: The Grammar of Visual Design*. London: Routledge.
- [7] Li Z. (2003). Social Semiotic Approach to Multimodal Discourse. *Foreign Languages Research*. 05:1-8+80. Doi:CNKI:SUN: NWYJ.0.2003-05-000.
- [8] Martinec, R. & Salway, A. (2005). A System for Image-Text Relation in New Media. *Visual Communication*, (3): 356-358.
- [9] Royce, T. (1998). *New Directions in the Analysis of Multimodal Discourse*. London: Lawrence Erlbaum Associates.
- [10] Zhu Y. (2007). "Theory and Methodology of Multimodal Discourse Analysis." *Foreign Language Research*. 05:82-86. doi:10.16263/j.cnki.23-1071/h.2007.05.034.
- [11] Zhang D. (2009). On A Synthetic Theoretical Framework for Multimodal Discourse Analysis [J]. *Foreign Languages in China*, 6 (01): 24-30. DOI:CNKI:SUN: ZGWE.0.2009-01-010.
- [12] Zhang D, and Yuan Y. (2011). On the Synergy of Different Modes in Dynamic Multimodal Discourse: Taking TV Weather Forecast Multimodal Discourse as an Example. *Shandong Foreign Language Teaching*, 2011, 32(05): 9-16. doi:10.16482/j.sdwy37-1026.2011.05.009.